Application of Bayesian Geostatistical Prediction to Estimate Local Variation in Numbers of People with Low and High Intensity *Schistosoma Haematobium* Infections in West Africa

Authors: Archie C A Clements[a,b], Sonja Firth[a], Robert Dembelé[c], Amadou Garba[d], Seydou Touré[e], Moussa Sacko[c], Aly Landouré[f], Elisa Bosqué-Oliva[g], Adrian G Barnett[h], Simon Brooker[i], Alan Fenwick[g]

[a]University of Queensland, School of Population Health, Herston, Queensland, Australia

[b]Australian Centre for International and Tropical Health, Queensland Institute of Medical Research, Herston, Queensland, Australia

[c]Programme National de Lutte Contre la Schistosomiase, Ministère de la Santé, Bamako, Mali

[d]Programme National de Lutte Contre la Bilharziose et les Géohelminthes, Ministère de la Santé Publique et de la Lutte Contre les Endémies, Niamey, Niger

[e]Programme National de Lutte Contre la Schistosomiase, Ministère de la Santé, Ouagadougou, Burkina Faso

[f]Institut National de Recherche en Santé Publique, Bamako, Mali

[g]Schistosomiasis Control Initiative, Imperial College, London, United Kingdom.

[h]Institute for Health and Biomedical Innovation, Queensland University of Technology, Kelvin Grove, Queensland, Australia.

[i]London School of Hygiene and Tropical Medicine, London, United Kingdom.

Corresponding author:

Archie Clements, School of Population Health, University of Queensland, Herston Road, Herston, 4006, Queensland, Australia. Ph: +61 (0) 7 33464706. Email a.clements@uq.edu.au

Abstract

Objective

We aimed to predict sub-national spatial variation in numbers of people infected with *Schistosoma haematobium*, and associated uncertainties, in Burkina Faso, Mali and Niger, prior to implementation of national control programmes.

Methods

We used national field survey datasets covering a contiguous area 2,750 × 850 km, from 26,790 school-aged children (5–14 years) in 418 schools. Bayesian geostatistical models were used to predict prevalence of high and low intensity infections and associated 95% credible intervals (CrI). Numbers infected were determined by multiplying predicted prevalence by numbers of school-aged children in 1 km^2 pixels covering the study area.

Findings

Numbers of school-aged children with low-intensity infections were: 433,268 in Burkina Faso, 872,328 in Mali and 580,286 in Niger. Numbers with high-intensity infections were: 416,009 in Burkina Faso, 511,845 in Mali and 254,150 in Niger. 95% CrIs (indicative of uncertainty) were wide; e.g. the mean number of boys aged 10–14 years infected in Mali was 140,200 (95% CrI 6200, 512,100).

Conclusion

National aggregate estimates for numbers infected mask important local variation, e.g. most *S. haematobium* infections in Niger occur in the Niger River valley. Prevalence of high-intensity infections was strongly clustered in foci in western and central Mali, north-eastern and north-western Burkina Faso and the Niger River valley in Niger. Populations in these foci are likely to carry the bulk of the urinary schistosomiasis burden and should receive priority for schistosomiasis control. Uncertainties in predicted prevalence and numbers infected should be acknowledged and taken into consideration by control programme planners.

Key words

Schistosomiasis; *Schistosoma haematobium*; Spatial analysis; Bayesian methods; Burden of disease

Introduction

Accurate estimation of the number of people in a population affected by a disease is important for correctly prioritizing control of that (versus another) disease and allocating adequate resources to control or prevention programmes. In the absence of high-quality surveillance data in developing countries, calculations of the magnitude of a disease problem often use prevalence estimates from cross-sectional surveys, which are rarely randomised or representative of the whole population. For schistosomiasis, caused by trematodes of the genus *Schistosoma*, a recent review reported a global burden of 207 million people infected and 779 million people at risk in 2003, the majority of whom reside in sub-Saharan Africa (SSA).[1] Prevalence estimates from a 1989 report[2] formed the basis of these calculations for most countries in SSA, highlighting the problem that prevalence data are often out of date. Additionally, estimates presented in the report were based on national aggregate data, while both the prevalence of schistosomiasis and populations at risk are known to be heterogeneously distributed within countries. Previous reports have ignored uncertainties in prevalence estimates and the size and spatial distribution of the population at-risk.

Recently, empirical maps of tropical infectious diseases have been used to improve estimates of populations infected and at-risk at the continental or global level[3-5] and increasingly to plan and target control programmes. Advances in the production of these maps include geostatistical prediction of the prevalence of infection with *S. haematobium* (the aetiological agent of urinary schistosomiasis),[6] other parasitic infections[7-10] and co-infections[11] using Bayesian methods.[12] The Bayesian approach is advantageous because covariate effects and spatial heterogeneity (clustering) can be modelled simultaneously, and predictions can be assessed for uncertainty.

While the planning of large-scale schistosomiasis control programmes and previous estimates of disease burden have used prevalence, it is well recognized that intensity of infection is a more informative metric for estimating morbidity (such as urinary tract lesions and anaemia)

associated with schistosome infections[13-15] and has a greater role in driving transmission than prevalence. Therefore, maps that represent the average intensity of infection, or distribution of low and high intensity infections, might provide more effective tools for burden of disease estimation and developing optimum intervention strategies than prevalence maps. An important statistical issue for intensity of parasitic infections is overdispersion (or aggregation), where most individuals have few parasites and few individuals have many parasites.[19] Spatial analysis[20-22] and prediction[23] of intensity of parasitic infections (*Wuchereria bancrofti* and *S. mansoni*) has been made using Bayesian models, with overdispersion in the individuals' parasite or egg counts modelled using the negative binomial distribution.

Burkina Faso, Mali and Niger, three contiguous countries in the Sahelian zone of West Africa, recently conducted coordinated national cross-sectional school-based parasitological surveys.[16] The surveys were unprecedented in their geographical range (approximately 2,750 km × 850 km), numbers of participants (26,790 school-aged children) and numbers of locations (418 schools). We aim to use the survey data to predict sub-national spatial distributions of prevalence of low and high-intensity infections of *S. haematobium* and use the prediction maps to calculate the numbers of school-aged children infected and at risk of disease. We also aim to estimate uncertainties in predicted prevalence and numbers infected.

Methods

Selection of schools and children

The SCI-supported programmes involve mass distribution of praziquantel (for urinary and intestinal schistosomiasis) and albendazole (for STH) as appropriate, but the parasitic infection with the highest prevalence in all three countries is *S. haematobium*. Therefore, the programmes were primarily designed to control urinary schistosomiasis [16] and here we only present analysis of the *S. haematobium* data. Ethical approvals for data collection were obtained from St. Mary's Hospital

Research Ethics Committee UK, the National Public Health Research Institute's (INRSP) scientific committee in Mali, the Ministry of Health Ethics and Scientific Committees in Burkina Faso and the Ministry of Health Ethical Committee in Niger.

Using sample size calculations based on historic data from Mali [17] and stratification according to the geographic area of the countries; it was decided to survey 87 schools in Burkina Faso, 226 schools in Mali and 215 schools in Niger, and 60 children per school. Ultimately, 418 of the 528 schools were surveyed as remote, sparsely-populated areas were excluded for logistical reasons. Different methods for spatial stratification to maximise geographical coverage were applied within the three countries. In Mali and Niger, sample frames were used that contained the location of all communities. Spatial stratification was done by overlaying a 1 decimal degree squared grid on these countries in a geographical information system (GIS; ArcView version 9, ESRI, Redlands, CA). Communities were selected from the cells using simple random selection and, where more than one school was present, a school was selected on arrival using simple random selection. In Burkina Faso, lists of schools were available for each province but they were not geo-referenced. The numbers of schools to be selected in each province was weighted according to the area of the province. School sampling was then done in each province using simple random selection.

The surveys were conducted from 2004–2006. On arrival, school coordinates were collected using a global positioning system and all available school-aged children were assembled. If there were less than 50 boys or girls, all individuals of that sex were selected (because of difficulties in compliance after excluding a minority of assembled individuals). If there were more than 50 boys or girls, 30 of that sex were selected using systematic random sampling. Urine and stool samples were collected from each child and processed using standard parasitological methods. A single 10 ml urine sample from each child was drawn into a syringe and passed through a filter, which was examined under a microscope in the field. Egg counts of *S. haematobium* were recorded and transferred to a Microsoft Access database.

Prediction of Prevalence of Low and High-Intensity Infections

Spatial prediction was based on Bayesian geostatistics.[18] Rather than modelling egg counts using the negative binomial distribution, we chose a multinomial model with egg counts categorised into: 1) no infection, 2) low-intensity infection (1–50 eggs/10ml urine) and 3) high-intensity infection (>50 eggs/10ml urine). This was for two reasons: 1) expediency, given that in Burkina Faso extremely high egg counts were recorded as >1000 eggs/10 ml urine, meaning that the upper tail of the distribution was truncated; and 2) to facilitate future estimation of the burden of schistosomiasis, where existing evidence for related morbidity is based on stratified egg counts, often using the World Health Organization definitions of low and high intensity.[24]

Age and sex have been demonstrated to be associated with urinary schistosomiasis prevalence, probably due to physiological differences in susceptibility.[25, 26] Distance from a perennial inland water body (PIWB) is a plausible risk factor for exposure and subsequent infection because transmission requires contact with aquatic habitats of the intermediate host snails, *Bulinus* spp. Distances were derived from an electronic PIWB map obtained from the Food and Agriculture Organisation. Effects of temperature and rainfall on the distribution of *Bulinus* snails are reviewed in Rollinson et al.[27] Satellite-derived mean land surface temperature (LST) and normalized difference vegetation index (NDVI; a proxy for rainfall) for 1982–1998 were obtained from the National Oceanographic and Atmospheric Administration's (NOAA) Advanced Very High Radiometer (AVHRR). The initial candidate set of variables included individual participant variables: sex and age (categorised into 5–9 and 10–14 years); and school-level ecological variables: distance from a PIWB, LST (with and without a quadratic term) and NDVI. We tested nominal and ordinal multinomial regression models and found a nominal model to provide a better fit. Variable selection was done using fixed-effects multinomial regression models in Stata/SE 10.0 (StataCorp, College Station, Texas). NDVI was excluded on the basis of a Wald's $P>0.2$ and all

remaining variables were selected for inclusion in the spatial model. This model and details on how prediction was done are presented in Box 1. Outputs of Bayesian models (including parameter estimates and spatial predictions) are termed posterior distributions, which fully represent uncertainties associated with their estimated values. We summarized the posterior distributions in terms of the posterior mean and the 95% credible interval (CrI) limits, between which the true value occurs with a probability of 95%.

Calculations of Numbers Infected

An electronic population surface for the study area was obtained from the Global Rural-Urban Mapping Project (GRUMP) alpha version[28] and imported into ArcView. GRUMP is a 30-arc second (1 km^2) population raster dataset that combines 2000 census data at a sub-national level with an urban extent mask. In GRUMP, the population is redistributed using an algorithm that assumes a greater proportion is located in urban areas.[29]

Country-specific population growth rates and proportions of the population by sex and age-group (5–9 and 10–14 years) from the United Nations Population Division – World Population Prospects[30] were used to project population surfaces to 2005. Surfaces representing the mean, lower and upper 95% CrI limits of predicted prevalence in each age-sex group were multiplied by the the 2005 projected population surfaces using the Spatial Analyst Extension of ArcView, giving numbers infected per 1 km^2 pixel, which were summed for each country.

Calculations of Numbers of People at Risk

By design, the multinomial model gave a predicted prevalence that was non-zero in all locations, even those where field data suggested no infections occurred. A receiver operating characteristic analysis was conducted to determine the optimal threshold (where sensitivity=specificity) of combined predicted prevalence (i.e. high plus low intensity infection prevalence) that best

discriminated between schools with zero and non-zero observed prevalence. Using this approach, a predicted prevalence threshold of 5.3% gave the best discriminatory performance. A mask, created in the GIS to exclude areas with combined predicted prevalence values less than or equal to 5.3%, was overlaid on the different population surfaces for each country to calculate numbers and proportions at risk of infection in the school-aged and total population.

Results

Prediction of Prevalence of Low and High-Intensity Infections

The total number of children aged 5–14 included in the surveys was 4,808 in Burkina Faso, 14,586 in Mali, and 7,396 in Niger. The raw prevalence of low intensity infections was 10.0% (95% confidence intervals 9.2, 10.9%) in Burkina Faso, 25.2% (24.5, 25.9%) in Mali and 10.1% (9.4, 10.8%) in Niger. For high intensity infections the raw prevalence was 8.5% (7.7, 9.3%) in Burkina Faso, 11.4% (10.9, 11.9%) in Mali and 3.4% (3.0, 3.8%) in Niger. A map of raw prevalence of *S. haematobium* infection is presented in Figure 1.

The spatial model is presented in Table 1: LST was not significantly associated with prevalence of low or high-intensity infections; distance from a PIWB was significantly and negatively associated with prevalence of both low and high-intensity infections; and boys and children aged 10–14 years had a significantly higher prevalence of low and high-intensity infections than girls or children aged 5–9 years. The rate of decay of spatial correlation was higher for low-intensity infections than high-intensity infections and the variance of the spatial random effect (the sill in geostatistical terms) was higher for high-intensity infections than low-intensity infections, indicating a stronger propensity for spatial clustering for high-intensity infections.

Separate prediction maps were produced for boys and girls aged 5–9 and 10–14. Illustrative examples of low and high-intensity infections in boys aged 10–14 (the highest prevalence group), are presented in Figures 2–3 (note, the maps for the other age-sex groups showed the same spatial

distribution but lower predicted prevalence, reflecting in the lower odds ratios for these groups). The spatial distribution of low and high-intensity infections were contrasting; low-intensity infections were more widespread but the variation in prevalence was less extreme, with predicted prevalence between 10 and 50% in most mapped areas of Burkina Faso (excluding the southwest), Mali (excluding the far south) and Niger (excluding central regions). High-intensity infections had a more restricted spatial distribution, with defined clusters of high (>50%) prevalence located in a mid-latitudinal band from western to central Mali, northern and central Burkina Faso and the Niger River valley in Niger. There were large delineated areas of low (<5%) prevalence in southern, northern and eastern Mali, south-western Burkina Faso and most of Niger, excluding the Niger River valley. The 95% CrI maps showed wide uncertainty (imprecision) in predicted prevalence estimates, though the prediction model was clearly able to exclude parts of each country as at-risk areas for significant transmission of *S. haematobium* (as indicated by upper 95% CrI limits of predicted prevalence that were <5%).

Numbers Infected and At Risk

The estimated numbers of school-aged children with low and high-intensity infections in the three countries, with associated 95% CrIs, are presented in Table 2. We estimated the numbers of school-age children with low-intensity infection, based on the means of the posterior distributions, to be 433,268 in Burkina Faso, 872,328 in Mali and 580,286 in Niger and the numbers with high-intensity infections to be 416,009 in Burkina Faso, 511,845 in Mali and 254,150 in Niger. The 95% CrIs for numbers infected in each age-sex group were very wide. Maps of the numbers of boys aged 10–14 years with high and low intensity infections (Figures 4–5) show that there was considerable within-country variation in the burden of schistosomiasis (this was apparent for boys and girls of both age groups – note that, as for predicted prevalence, the spatial patterns were identical for each age-sex group, but the proportion infected was uniformly lower for girls and older boys).

We estimated that 3.5, 3.4 and 2.8 million school age children and a total of 12.5 million (94.5% of the total), 11.8 million (87.6% of the total) and 9.9 million (70.3% of the total) individuals of all ages were at-risk of urinary schistosomiasis in Burkina Faso, Mali and Niger, respectively. The maps confirm that there were parts of each country where people were not at-risk and that could be potentially excluded from active surveillance or nationally coordinated schistosomiasis control.

Discussion

In this report we used robust, contemporary statistical methods for spatial prediction in a novel application, that is, to estimate local heterogeneity in the distribution of high and low intensity parasitic infections for disease burden estimation, in the context of control programme planning. The numbers of people estimated by Steinmann et al.[1] to be infected with *Schistosoma* spp. in 2003 were 7.8 million in Burkina Faso, 7.8 million in Mali and 3.2 million in Niger, based on prevalence estimates of 60.0%, 60.0% and 26.7% and populations at risk of 13.0, 13.0 and 12.0 million (100% of the estimated population of each country). If we assume, as Steinmann *et al*. did, that prevalence is the same for all age groups (which overestimates prevalence given that school-age children are usually the highest-prevalence group in the population[31]), our calculations of the total number infected with *S. haematobium* are 3.0 million (23.0% of the total) in Burkina Faso, 4.8 million (35.4% of the total) in Mali and 3.0 million (21.1% of the total) in Niger. Based on these differences we conclude that the previously reported numbers infected were considerably overestimated for Burkina Faso and Mali and the numbers at-risk were over-estimated for all three countries. Our confidence in this conclusion is based on the fact that our data were recent, extensive, randomised and representative. Our calculations excluded *S. mansoni* infections, which were included in the figures of Steinmann *et al*., but which occurred in our surveys with very a low prevalence of 0.5% and 0.3% in Burkina Faso and Niger, and moderately low prevalence of 6.7% in Mali. More

importantly, the figures reported by Steinmann *et al*. overlook important heterogeneities in the spatial distribution of people infected and at-risk of schistosomiasis that are captured in our maps.

We have represented uncertainties in the numbers infected with *S. haematobium*, both at the national and sub-national level, which were wide even with the large sample size and geographical coverage of the data available. It might be suggested that the large uncertainties demonstrated in our prediction maps limit their utility for decision making. However, we argue that, given tools now exist for honest and accurate uncertainty representation in spatial predictions and the continued need for disease maps in planning control programmes, it is beneficial to be aware of, acknowledge and take into consideration such uncertainties when interpreting maps for disease control. The true level of uncertainty will be even greater because we did not include uncertainties in the projected population size or population migration between areas, which are considerable in these three countries. GRUMP, and GPW v3 on which it is based, use a variety of data sources and methods which can result in national population totals that do not match UN totals. In 2000, GPW v3 population estimates for both Mali and Niger differed from UN estimates by more than 5%,[28] increasing uncertainty in our updated 2005 population estimates for these countries. We also assumed an even growth nationally, which may underestimate the variability in population growth across different regions and age-groups. The limitations of ancillary data used for GRUMP calculations have been noted[29] but it is currently impossible to determine, for different parts of the study area, the direction (over or under-estimation) or magnitude of inaccuracy caused by these limitations.

Our surveys were spatially stratified, which ensured equal geographical coverage in high- and low-population-density areas, and more even precision of predicted prevalence estimates throughout the study area, but heavier representation of less densely populated areas than would have been the case using a non-stratified approach. Because prevalence is likely to be different in low- and high-density areas, sample weighting is necessary to ensure accurate national prevalence

estimates; our map calculations are essentially a sample weighting approach. This is evident from the contrast between raw prevalence estimates for low and high-intensity infections combined of 18.5% for Burkina Faso, 36.6% for Mali and 13.5% for Niger, and the map-derived estimates, which were 23.0%, 35.4% and 21.1% respectively. The most striking difference was for Niger; this arose because the highest prevalence area was the Niger River valley, also the area with the highest population density, which was under represented (from a population perspective) in the survey relative to low-density areas in the north and east.

The different spatial distributions of low and high-intensity infections are noteworthy and in agreement with Guyatt et al.[19] who demonstrated that the relationship between overall prevalence and prevalence of high intensity infections varied geographically in different parts of Africa. While low-intensity infections were more widespread, there was less spatial variability than high-intensity infections, which demonstrated greater aggregation within clusters.[19] Future control programmes will have the greatest impact on morbidity and transmission if they focus on these high-intensity infection clusters. It should also be noted that the sensitivity and specificity of microscopic diagnosis is likely to differ between high and low-intensity infections, potentially influencing the different spatial distributions and statistical estimates of spatial heterogeneity. Quantifying these influences is an area of future research.

The maps presented in this report are currently being used by national programme managers as objective decision-support tools for geographically targeting existing resources more efficiently to high-risk communities. They also have several other potential uses. First, they can be used by national programme managers to advocate for appropriate resources from governmental or international donor sources (particularly after SCI funding ends). Second, they can be used to formulate and compare different control strategies to determine the likely impact on transmission and morbidity. And lastly, they can be used for advocacy and empowerment at the sub-national level, whereby local resource needs and priorities for schistosomiasis control – often subsumed in

aggregate national data – can be presented to national programme managers and government officials. We encourage national programme managers in other countries, and those who focus on other diseases, to conduct spatially stratified disease surveys and undertake mapping of sub-national disease distributions to provide evidence for more efficient targeting of their resources.

Acknowledgements

We thank the children, parents, teachers and head teachers who participated in the surveys. We also thank the technicians and support staff who undertook the surveys and members of the national schistosomiasis control programmes who provided administrative and organisational assistance. At the time of this study, SCI was supported by the Bill and Melinda Gates Foundation. SB is supported by a Career Development Fellowship (081673) from the Wellcome Trust.

Box 1. The Bayesian multinomial regression model with geostatistical random effects.

The spatial model was fitted in WinBUGS version 1.4 (MRC Biostatistics Unit, Cambridge, and Imperial College London, UK). The individual data were aggregated into age and sex groups and by location. Using three infection outcome groups (1 = no infection, 2 = low intensity infection, 3 = high intensity infection), we assumed

$$Y_{ijk} \sim \text{Multinomial}(p_{ijk}, n_{ijk}),$$

$$p_{ijk} = \frac{\phi_{ijk}}{\sum_{k} \phi_{ijk}},$$

where Y_{ijk} is the observed number positive, n_{ijk} is the number tested and p_{ijk} is probability of infection at location i, in age-sex group j, infection outcome group k. To give a reference value, ϕ_{ij1} was constrained to equal one. For the other outcome groups we fitted the nominal regression models

$$log(\phi_{ij2}) = \alpha_2 + \sum_{z=1}^{T} \beta_{z2} \times x_{zij} + \theta_{i2},$$

$$log(\phi_{ij3}) = \alpha_3 + \sum_{z=1}^{T} \beta_{z3} \times x_{zij} + \theta_{i3},$$

where α_k is the outcome group-specific intercept, x is a matrix of covariates and β is a matrix of coefficients. The θ_{ik} are geostatistical random effects for prevalence of low and high-intensity infections defined by isotropic powered exponential spatial correlation functions

$$f(d_{ab}; \phi) = \exp[-(\phi d_{ab})],$$

where d_{ab} are the distances between pairs of points a and b, and ϕ is the rate of decline of spatial correlation per unit of distance. Non-informative priors were used for the intercepts (uniform prior with bounds $-\infty$ and ∞) and the coefficients (normal prior with mean = 0 and precision = 1×10^{-4}). The prior distribution of ϕ was also uniform with upper and lower bounds set at 0.06 and 50. The precisions of θ_{ik} were given non-informative gamma distributions.

A burn-in of 4,000 Markov chain Monte Carlo iterations was used, followed by 1,000 iterations where values for the intercept and coefficients were stored. Diagnostic tests for convergence of the stored variables were undertaken, including visual examination of history and density plots of the three chains. Convergence was successfully achieved after 5,000 iterations and the model was run for a further 10,000 iterations, during which predicted prevalence at the prediction locations was stored for each age and sex group.

Predictions of the prevalence of low and high intensity infections were made at the nodes of a 0.15×0.15 decimal degree (approximately 18 km^2) grid in WinBUGS using spatial model and the

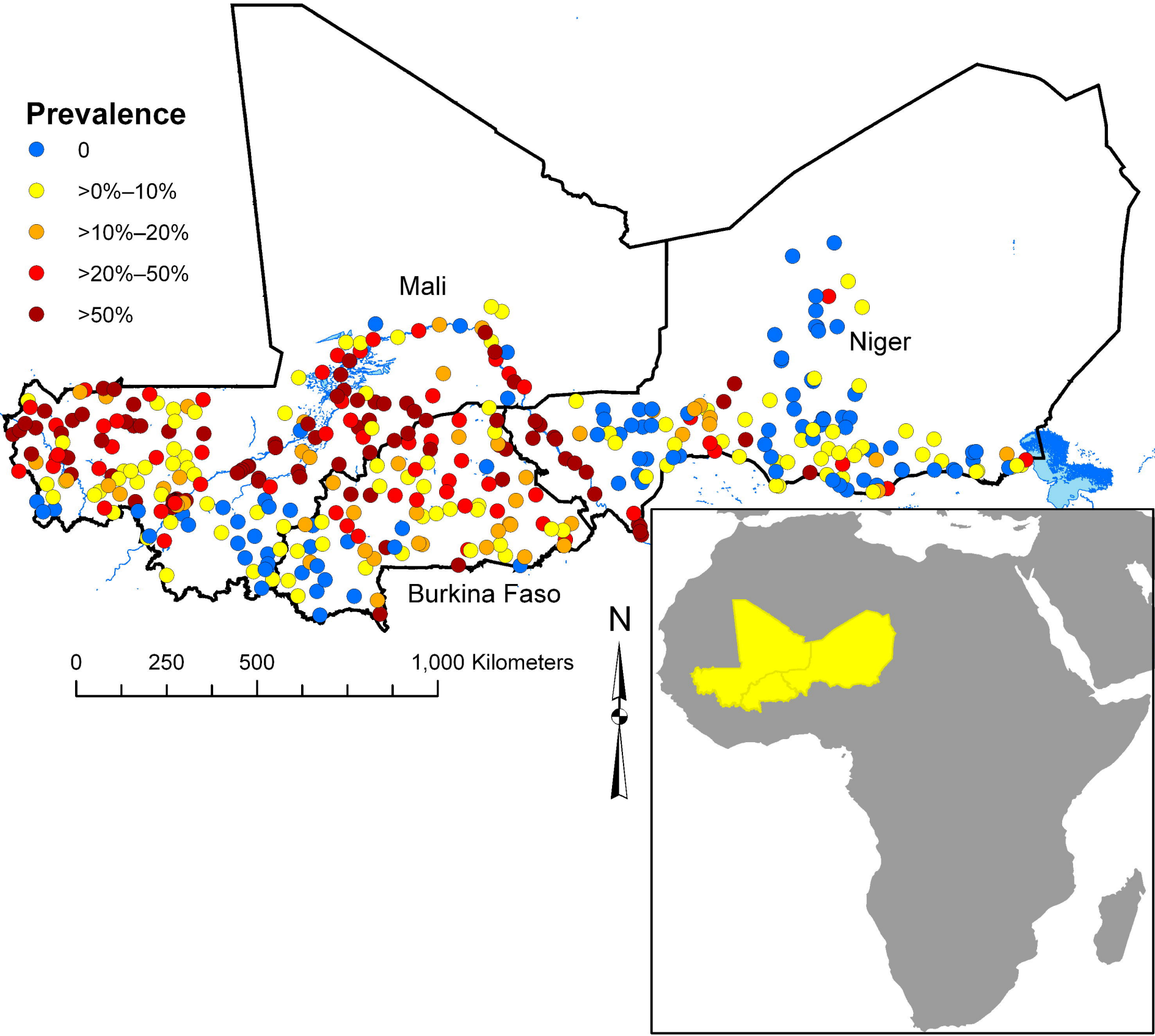
spatial.unipred command, which interpolated (using kriging) the spatial random effects for low and high-intensity infections. Predicted prevalences were calculated by adding the interpolated random effect to the sum of the products of the coefficients for the fixed effects and the values of the fixed effects at each prediction location. For the individual-level fixed effects (sex and age), separate calculations were done, where the coefficient for the relevant age and sex group were added to the sum. The overall sum was then back-transformed from the logit scale to the prevalence scale, giving prediction surfaces for prevalence of low and high-intensity infection in each age and sex group.

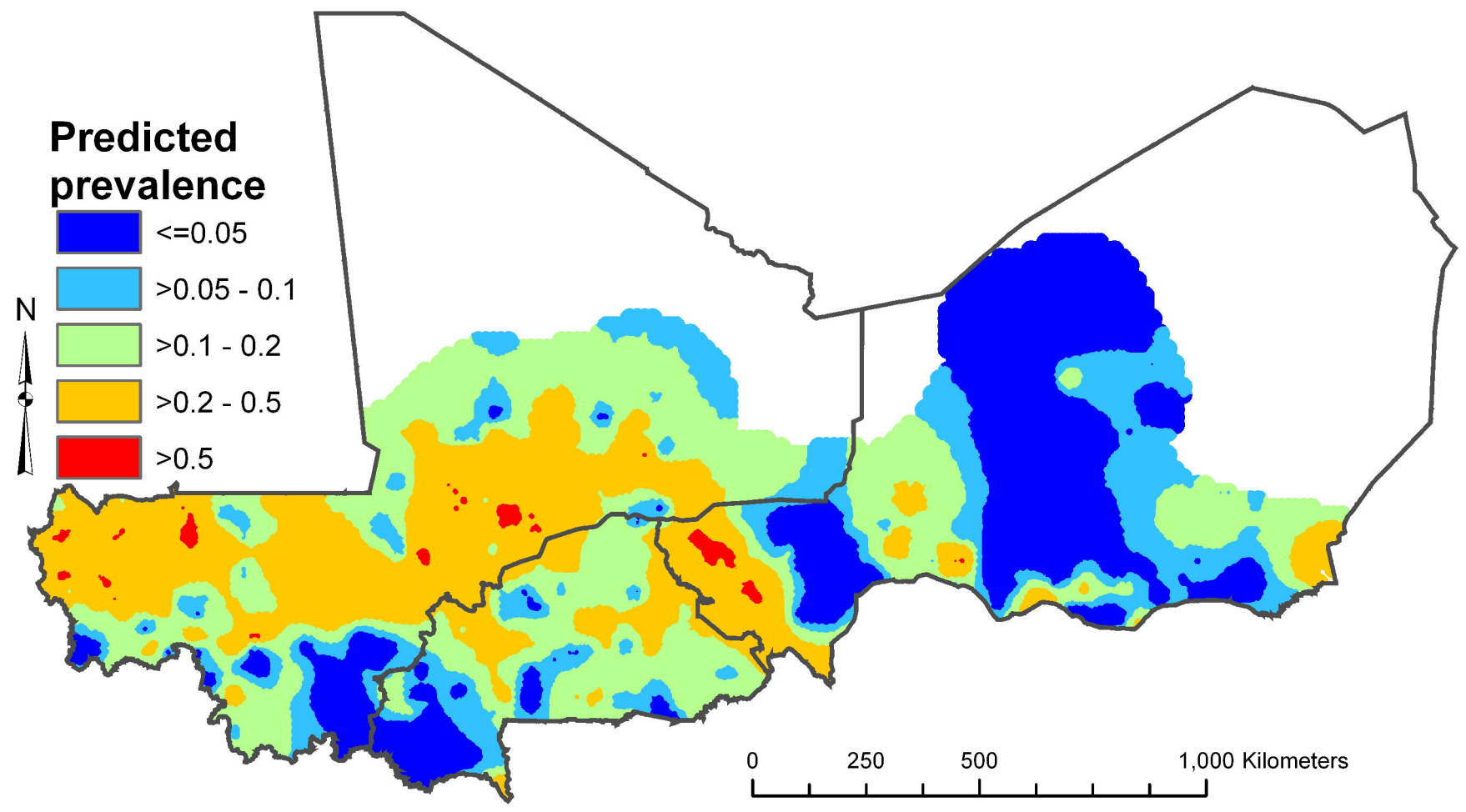
List of Figures:

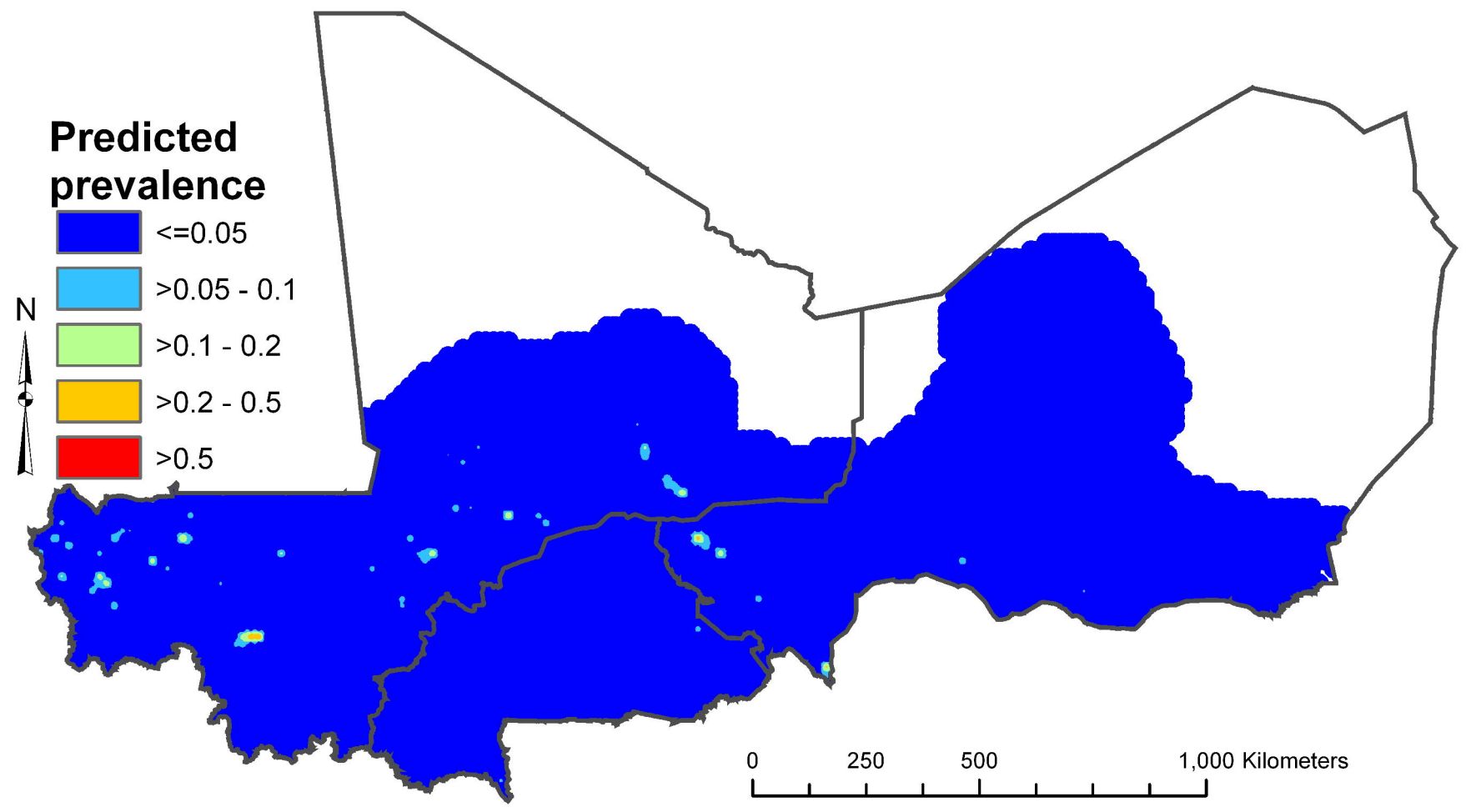
References

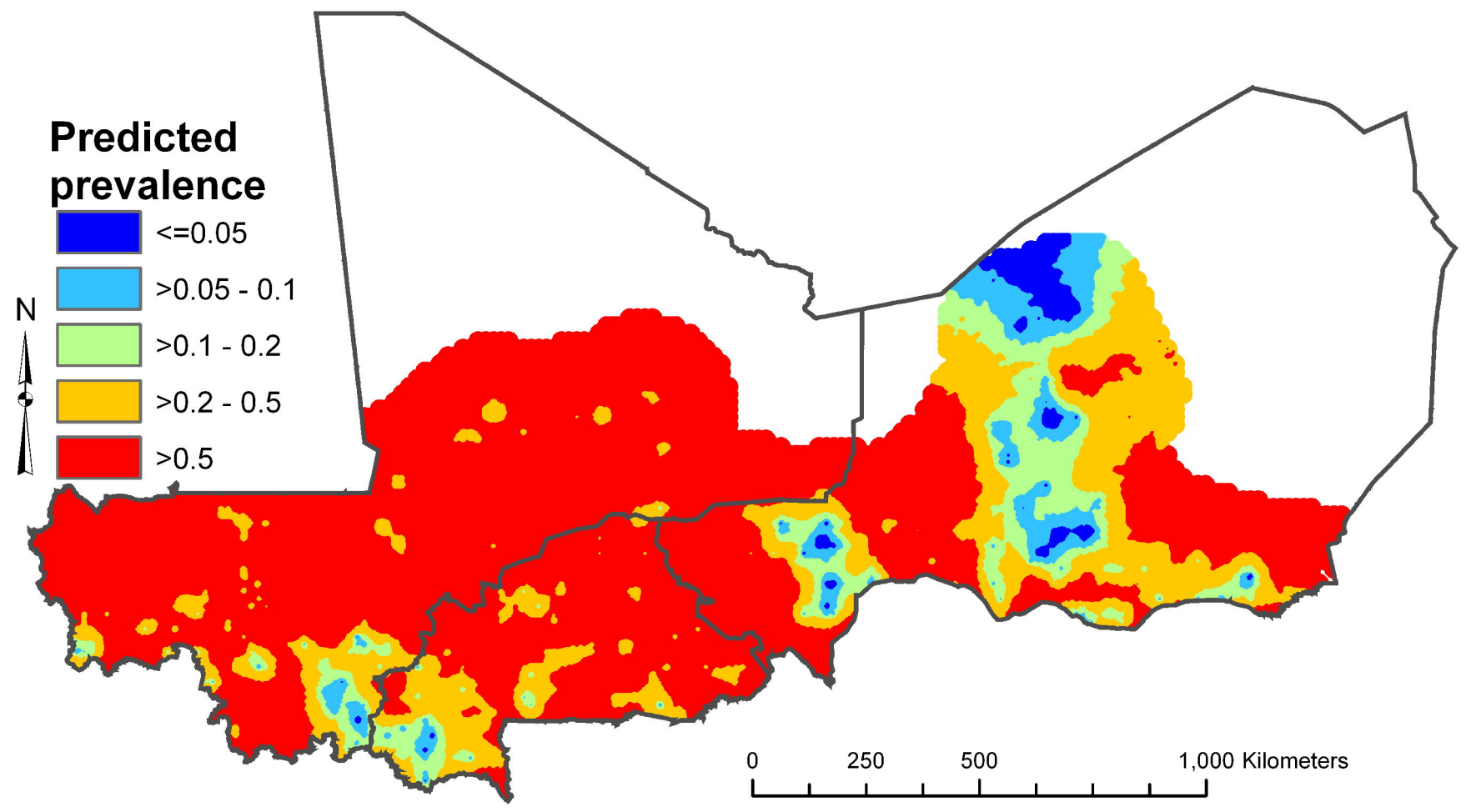
1. Steinmann P, Keiser J, Bos R, Tanner M, Utzinger J. Schistosomiasis and water resources development: systematic review, meta-analysis, and estimates of people at risk. *Lancet Infect Dis.* 2006;6(7):411-425.
2. Utroska JA, Chen MG, Dixon H, Yoon S, Helling-Borda M, Hogerzeil HV, et al. *An estimate of global needs for praziquantel within schistosomiasis control programmes.* Geneva: World Health Organization; 1989.
3. Brooker S, Clements AC, Bundy DA. Global epidemiology, ecology and control of soil-transmitted helminth infections. *Adv Parasitol.* 2006;62:221-261.
4. Guerra CA, Snow RW, Hay SI. Mapping the global extent of malaria in 2005. *Trends Parasitol.* 2006;22(8):353-358.
5. Hay SI, Guerra CA, Tatem AJ, Noor AM, Snow RW. The global distribution and population at risk of malaria: past, present, and future. *Lancet Infect Dis.* 2004;4(6):327-336.
6. Clements AC, Lwambo NJ, Blair L, Nyandindi U, Kaatano G, Kinung'hi S, et al. Bayesian spatial analysis and disease mapping: tools to enhance planning and implementation of a schistosomiasis control programme in Tanzania. *Trop Med Int Health.* 2006;11(4):490-503.
7. Gemperli A, Sogoba N, Fondjo E, Mabaso M, Bagayoko M, Briet OJ, et al. Mapping malaria transmission in West and Central Africa. *Trop Med Int Health.* 2006;11(7):1032-1046.
8. Gemperli A, Vounatsou P, Kleinschmidt I, Bagayoko M, Lengeler C, Smith T. Spatial patterns of infant mortality in Mali: the effect of malaria endemicity. *Am J Epidemiol.* 2004;159(1):64-72.
9. Raso G, Matthys B, N'Goran EK, Tanner M, Vounatsou P, Utzinger J. Spatial risk prediction and mapping of Schistosoma mansoni infections among schoolchildren living in western Cote d'Ivoire. *Parasitology.* 2005;131(Pt 1):97-108.
10. Raso G, Vounatsou P, Gosoniu L, Tanner M, N'Goran EK, Utzinger J. Risk factors and spatial patterns of hookworm infection among schoolchildren in a rural area of western Cote d'Ivoire. *Int J Parasitol.* 2006;36(2):201-210.
11. Raso G, Vounatsou P, Singer BH, N'Goran EK, Tanner M, Utzinger J. An integrated approach for risk profiling and spatial prediction of Schistosoma mansoni-hookworm coinfection. *Proc Natl Acad Sci U S A.* 2006;103(18):6934-6939.
12. Diggle P, Moyeed R, Rowlingson B, Thompson M. Childhood malaria in the Gambia: a case-study in model-based geostatistics. *Applied Statistics.* 2002;51:493-506.
13. Vester U, Kardorff R, Traore M, Traore HA, Fongoro S, Juchem C, et al. Urinary tract morbidity due to Schistosoma haematobium infection in Mali. *Kidney international.* 1997;52(2):478-481.
14. Abdel-Wahab MF, Esmat G, Ramzy I, Fouad R, Abdel-Rahman M, Yosery A, et al. Schistosoma haematobium infection in Egyptian schoolchildren: demonstration of both hepatic and urinary tract morbidity by ultrasonography. *Trans R Soc Trop Med Hyg.* 1992;86(4):406-409.
15. Warren KS, Mahmoud AA, Muruka JF, Whittaker LR, Ouma JH, Arap Siongok TK. Schistosomiasis haematobia in coast province Kenya. Relationship between egg output and morbidity. *Am J Trop Med Hyg.* 1979;28(5):864-870.
16. Garba A, Toure S, Dembele R, Bosque-Oliva E, Fenwick A. Implementation of national schistosomiasis control programmes in West Africa. *Trends Parasitol.* 2006;22(7):322-326.
17. Gryseels B. *Mission d'Evaluation du report, Lutte contre la schistosomiasis au Mali.* Eschborn: Gesellschaft für Technische Zusammenarbeit 1989.
18. Diggle P, Tawn J, Moyeed R. Model-based geostatistics. *Applied Statistics.* 1998;47:299–350.

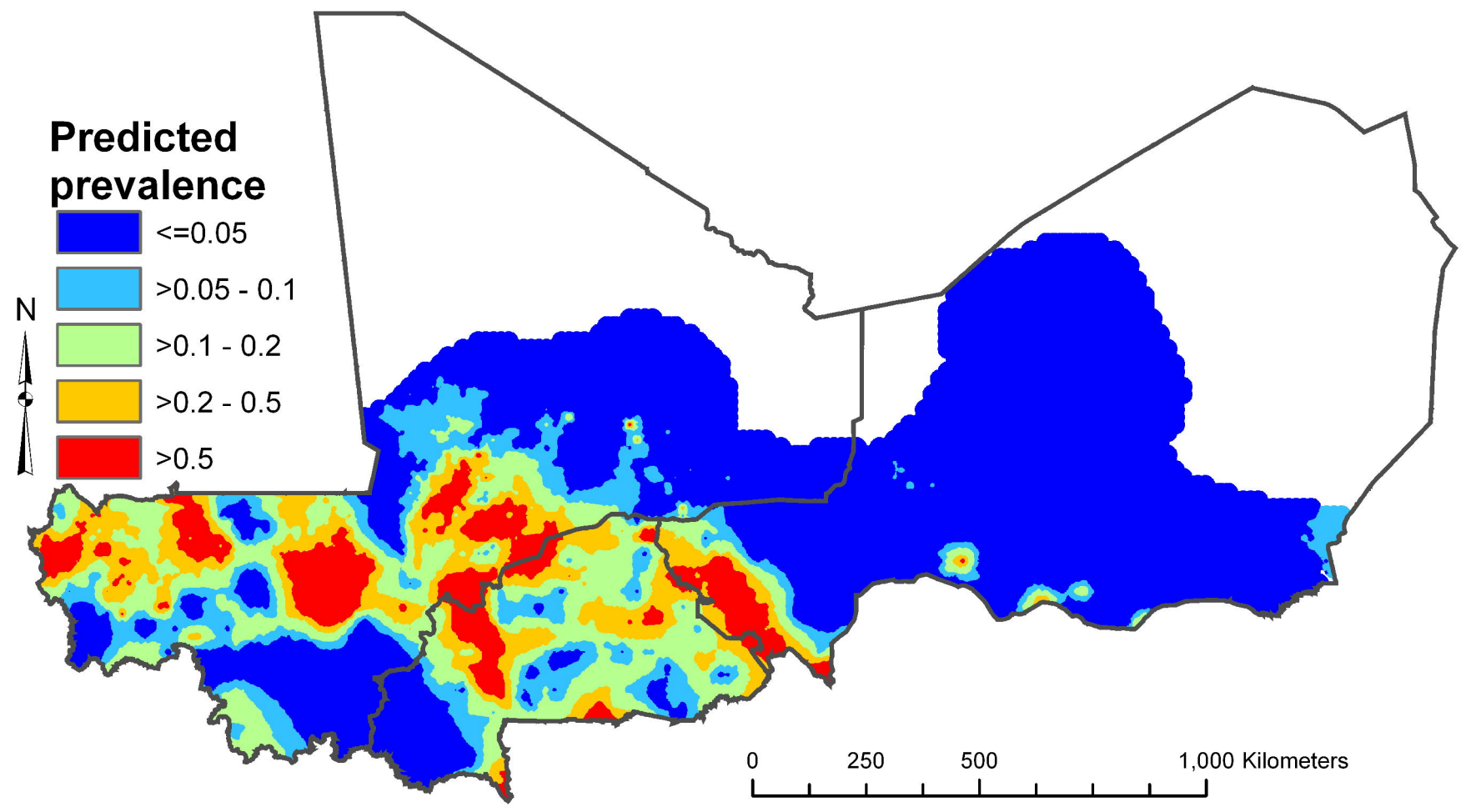
19. Guyatt HL, Smith T, Gryseels B, Lengeler C, Mshinda H, Siziya S, et al. Aggregation in schistosomiasis: comparison of the relationships between prevalence and intensity in different endemic areas. *Parasitology.* 1994;109 (Pt 1):45-55.
20. Alexander N, Moyeed R, Stander J. Spatial modelling of individual-level parasite counts using the negative binomial distribution. *Biostatistics (Oxford, England).* 2000;1(4):453-463.
21. Alexander ND, Moyeed RA, Hyun PJ, Dimber ZB, Bockarie MJ, Stander J, et al. Spatial variation of Anopheles-transmitted Wuchereria bancrofti and Plasmodium falciparum infection densities in Papua New Guinea. *Filaria journal.* 2003;2(1):14.
22. Brooker S, Alexander N, Geiger S, Moyeed RA, Stander J, Fleming F, et al. Contrasting patterns in the small-scale heterogeneity of human helminth infections in urban and rural environments in Brazil. *Int J Parasitol.* 2006;36(10-11):1143-1151.
23. Clements AC, Moyeed R, Brooker S. Bayesian geostatistical prediction of the intensity of infection with Schistosoma mansoni in East Africa. *Parasitology.* 2006;133(Pt 6):711-719.
24. WHO. *Coordinated use of anthelminthic drugs in control interventions: a manual for health professionals and programme managers.* Geneva: World Health Organisation; 2006.
25. Fulford AJ, Webster M, Ouma JH, Kimani G, Dunne DW. Puberty and Age-related Changes in Susceptibility to Schistosome Infection. *Parasitology today (Personal ed.* 1998;14(1):23-26.
26. Feldmeier H, Poggensee G, Krantz I. Puberty and Age-intensity Profiles in Schistosome Infections: Another Hypothesis. *Parasitology today (Personal ed.* 1998;14(10):435.
27. Rollinson D, Stothard JR, Southgate VR. Interactions between intermediate snail hosts of the genus Bulinus and schistosomes of the Schistosoma haematobium group. *Parasitology.* 2001;123 Suppl:S245-260.
28. Global Rural-Urban Mapping Project (GRUMP), Alpha Version. Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University. Available at http://sedac.ciesin.columbia.edu/gpw.
29. Balk DL, Deichmann U, Yetman G, Pozzi F, Hay SI, Nelson A. Determining global population distribution: methods, applications and data. *Adv Parasitol.* 2006;62:119-156.
30. World population prospects: the 2006 revision. United Nations, Department of Economic and Social Affairs, Population Division. Available at http://www.un.org/esa/population/publications/wpp2006/wpp2006.htm.
31. Brooker S, Donnelly CA, Guyatt HL. Estimating the number of helminthic infections in the Republic of Cameroon from data on infection prevalence in schoolchildren. *Bull World Health Organ.* 2000;78(12):1456-1465.

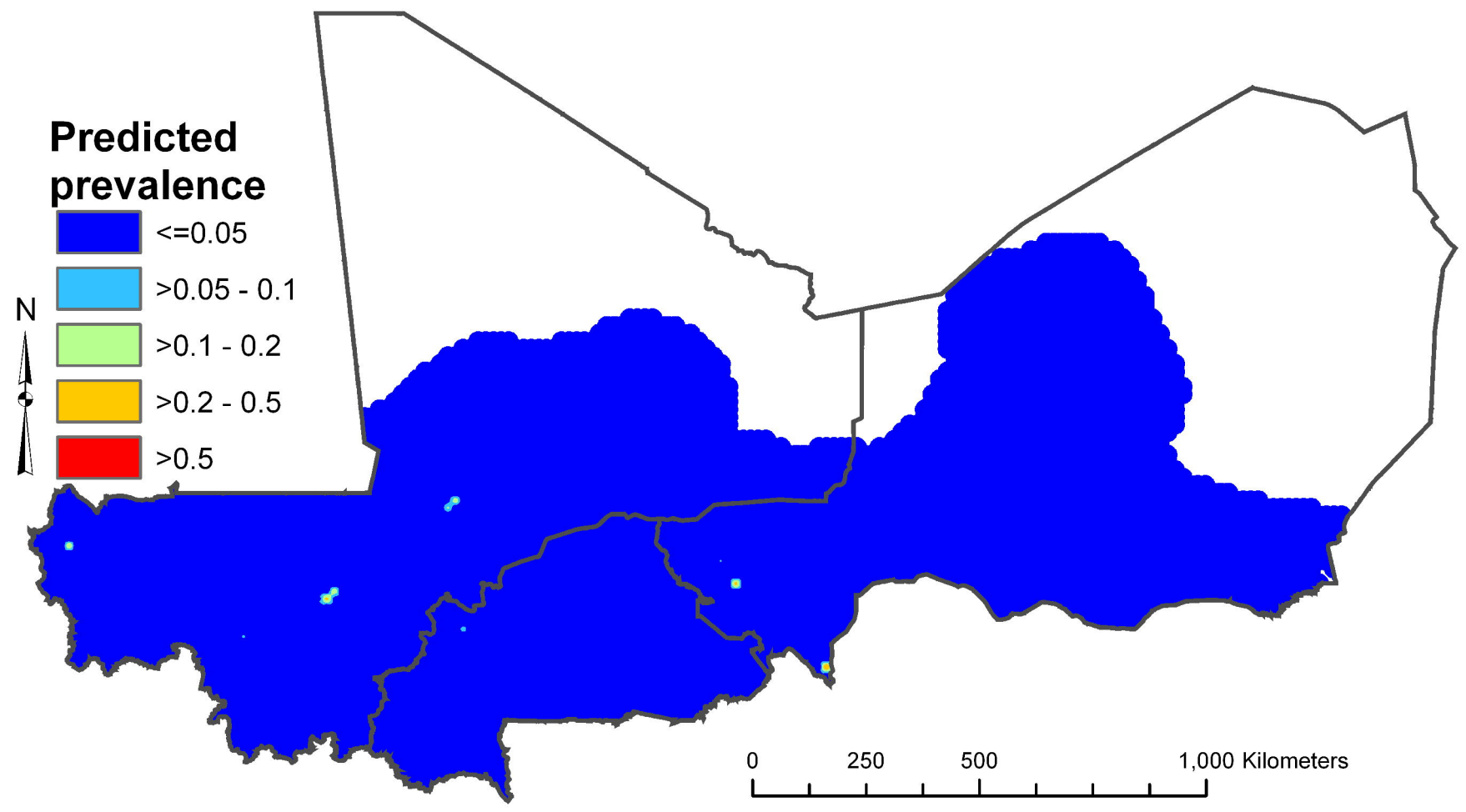

Prevalence
0
>0%–10%
>10%–20%
>20%–50%
>50%
Mali
Niger
Burkina Faso
0
250
500
1,000 Kilometers
N

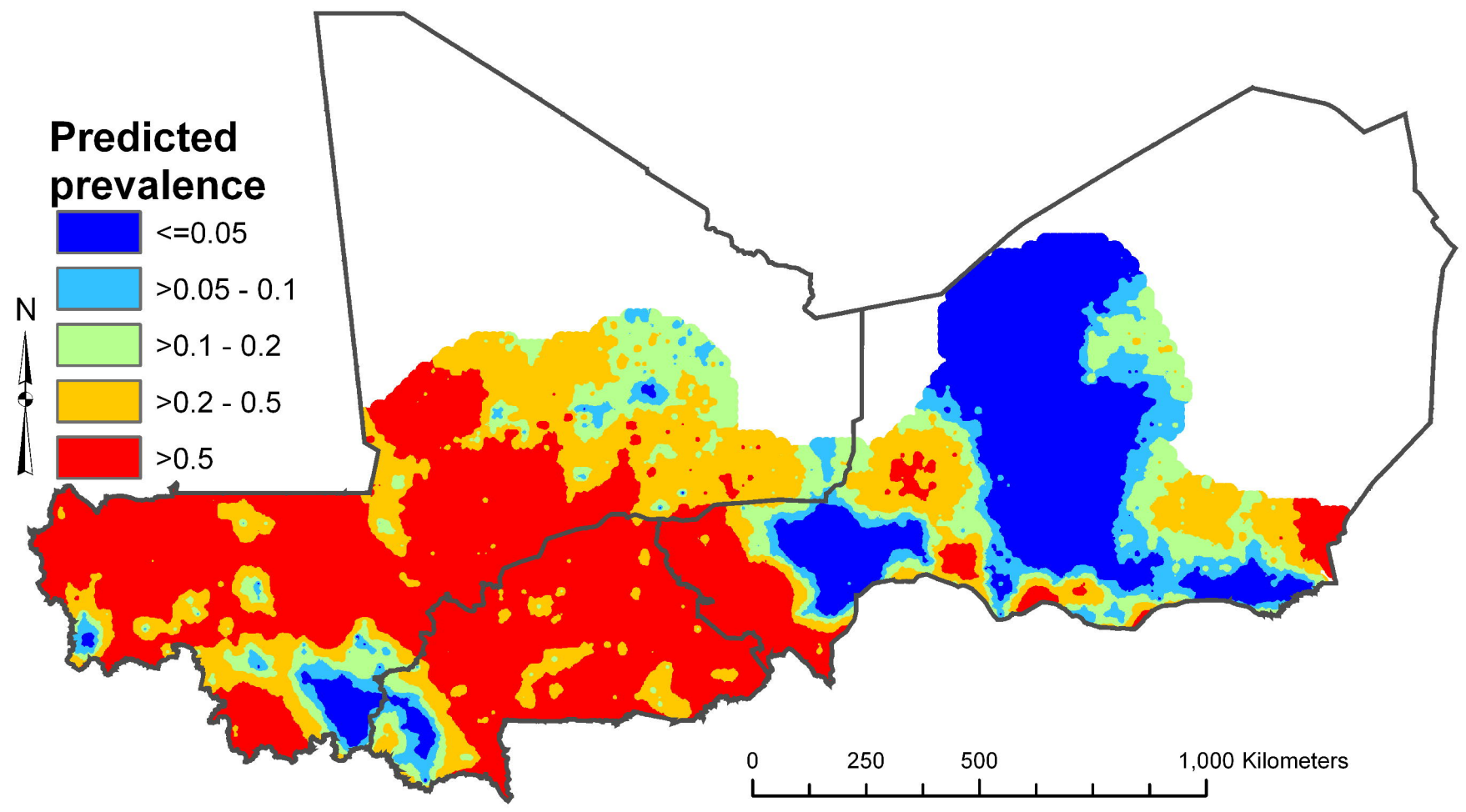

Predicted prevalence
<=0.05
>0.05 - 0.1
>0.1 - 0.2
>0.2 - 0.5
>0.5
N
0
250
500
1,000 Kilometers

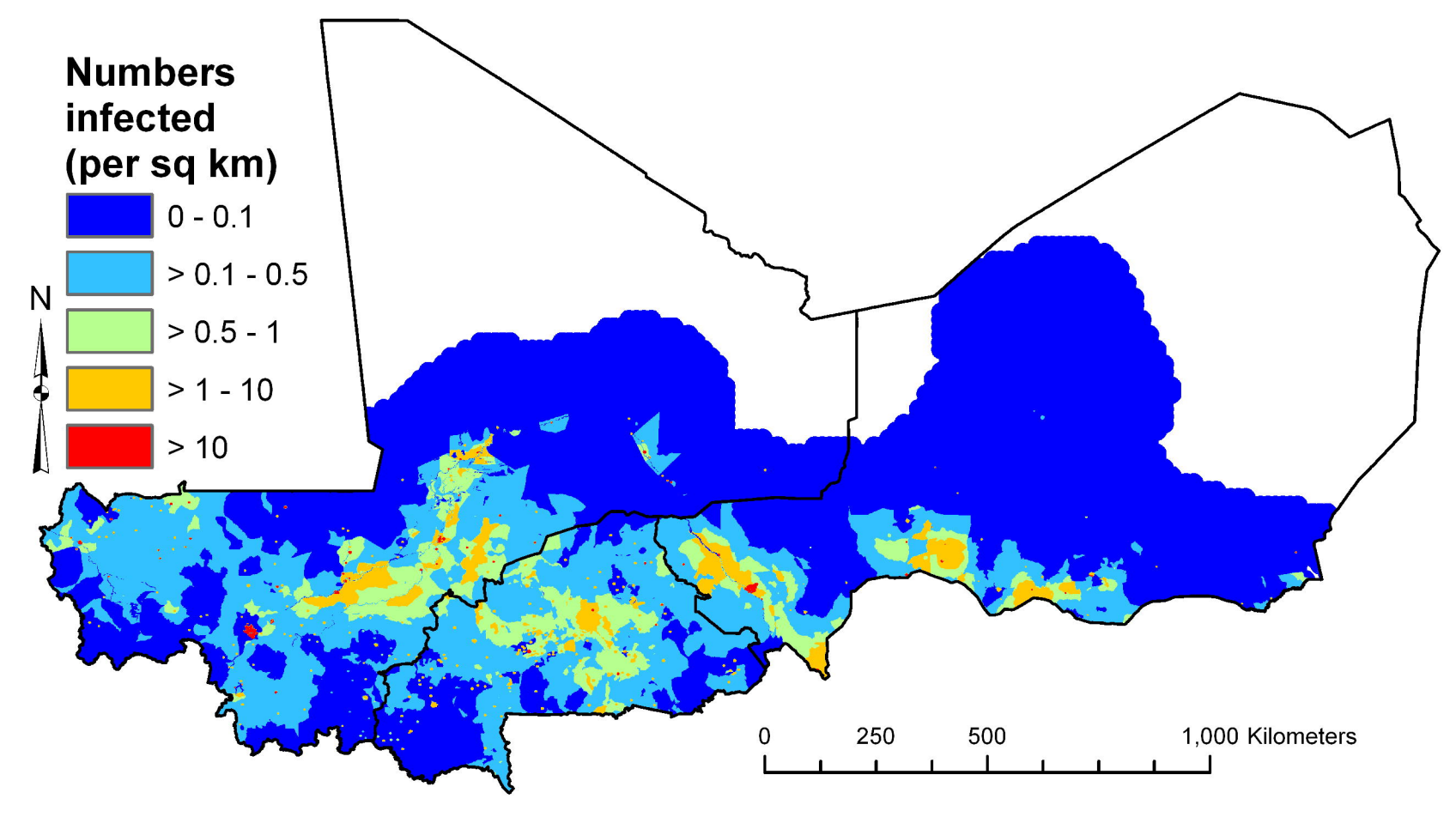

Predicted prevalence
<=0.05
>0.05 - 0.1
>0.1 - 0.2
>0.2 - 0.5
>0.5
N
0
250
500
1,000 Kilometers


Predicted prevalence
<=0.05
>0.05 - 0.1
>0.1 - 0.2
>0.2 - 0.5
>0.5
N
0
250
500
1,000 Kilometers


Predicted prevalence
<=0.05
>0.05 - 0.1
>0.1 - 0.2
>0.2 - 0.5
>0.5
N
0
250
500
1,000 Kilometers


Predicted prevalence
<=0.05
>0.05 - 0.1
>0.1 - 0.2
>0.2 - 0.5
>0.5
N
0
250
500
1,000 Kilometers


Predicted prevalence
<=0.05
>0.05 - 0.1
>0.1 - 0.2
>0.2 - 0.5
>0.5
N
0
250
500
1,000 Kilometers


Numbers infected (per sq km)
0 - 0.1
> 0.1 - 0.5
> 0.5 - 1
> 1 - 10
> 10
N
0
250
500
1,000 Kilometers

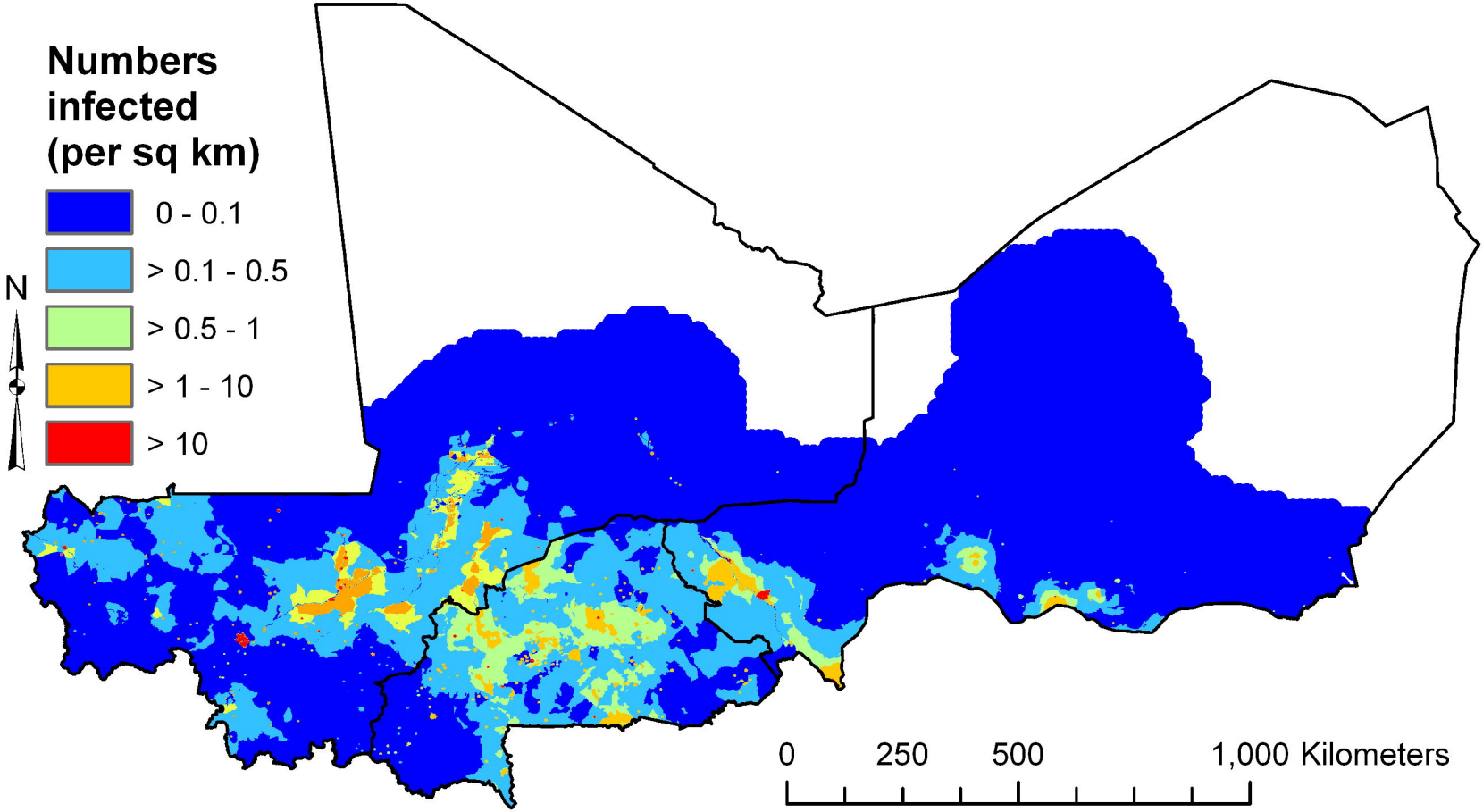

Numbers infected (per sq km)
0 - 0.1
> 0.1 - 0.5
> 0.5 - 1
> 1 - 10
> 10
N
0
250
500
1,000 Kilometers

Table 1. Bayesian multinomial regression model for low (1–50 eggs/10 ml urine) and high (>50 eggs/10 ml urine) intensity *Schistosoma haematobium* infections compared to no infection, in school-aged children, Burkina Faso, Mali and Niger, 2004–2006.

Variable	Low intensity infections Mean (95% CrI)	High intensity infections Mean (95% CrI)
Covariates (odds ratios):		
Land surface temperature	0.87 (0.65, 1.24)	0.34 (0.12, 0.57)
Land surface temperature^2	1.11 (0.94, 1.34)	1.00 (0.78, 1.30)
Distance from a PIWB	0.36 (0.24, 0.50)	0.24 (0.11, 0.52)
Age: 10–14 years (vs 5–9 years)	1.44 (1.31, 1.58)	1.45 (1.25, 1.65)
Sex: Female (vs Male)	0.80 (0.74, 0.86)	0.49 (0.44, 0.54)
Other model parameters:		
Intercept (log odds scale)	–3.04 (–3.62, –2.51)	–4.60 (–5.13, –3.39)
ϕ (rate of decay of spatial correlation)	2.03 (1.37, 2.82)	1.49 (0.75, 2.29)
σ^2 (variance of spatial random effect; sill)	5.80 (4.53, 7.94)	12.75 (8.92, 22.27)

PIWB = Perennial Inland Water-Body. CrI = Bayesian credible interval. Scales of odds ratios were standardised for land surface temperature and distance from a PIWB (data were orthogonalized to have a mean of zero and standard deviation of one).

Table 2. Estimated numbers of school-aged children (1000s) with low and high-intensity *S. haematobium* infections in Burkina Faso, Mali and Niger, using a Bayesian geostatistical model based on national survey data, 2004–2006.

Age-sex group	Burkina Faso		Mali		Niger	
High intensity infections	Pop.	Number infected Mean (95% CrI)	Pop.	Number infected Mean (95% CrI)	Pop.	Number infected Mean (95% CrI)
Boys: 5–9 y	1011.0	123.9 (1.4, 603.5)	1063	150.6 (6.6, 567.7)	1119.9	79.3 (1.6, 342.6)
Boys: 10–14 y	867.2	124.0 (1.5, 563.0)	896.5	140.2 (6.2, 512.2)	905.9	70.8 (1.4, 303.1)
Girls: 5–9 y	976.8	83.5 (0.8, 463.0)	1057.2	114.3 (4.1, 462.2)	1065.9	55.0 (1.0, 257.7)
Girls: 10–14 y	840.9	84.6 (0.9, 441.9)	895.3	106.8 (3.9, 421.2)	860.3	49.1 (0.9, 226.1)
Low intensity infections	Pop.	Number infected Mean (95% CrI)	Pop.	Number infected Mean (95% CrI)	Pop.	Number infected Mean (95% CrI)
Boys: 5–9 y	1011.0	112.0 (2.3, 528.2)	1063	226.9 (18.4, 687.8)	1119.9	157.4 (6.0, 543.9)
Boys: 10–14 y	867.2	115.9 (2.5, 508.3)	896.5	218.2 (19.0, 619.8)	905.9	146.2 (5.6, 486.3)
Girls: 5–9 y	976.8	100.3 (2.3, 481.3)	1057.2	216.6 (18.5, 659.9)	1065.9	143.4 (7.2, 485.7)
Girls: 10–14 y	840.9	105.1 (2.5, 468.1)	895.3	210.7 (19.6, 599.5)	860.3	133.2 (6.8, 434.2)

Pop = estimated population; mean = mean of posterior distribution of Bayesian estimates;
CrI = Bayesian credible interval.